Validation of diffusion MRI phenotypes for predicting response to bevacizumab in recurrent glioblastoma: post-hoc analysis of the EORTC-26101 trial.

Schell, Marianne ; Pflüger, Irada ; Brugnara, Gianluca ; Isensee, Fabian ; Neuberger, Ulf ; Foltyn, Martha ; Kessler, Tobias ; Sahm, Felix ; Wick, Antje ; Nowosielski, Martha ; Heiland, Sabine ; Weller, Michael ; Platten, Michael ; Maier-Hein, Klaus H ; von Deimling, Andreas ; van den Bent, Martin J ; Gorlia, Thierry ; Wick, Wolfgang ; Bendszus, Martin ; Kickingereder, Philipp

Abstract: BACKGROUND: This study validated a previously described diffusion-MRI phenotype as a potential predictive imaging biomarker in patients with recurrent glioblastoma receiving bevacizumab (BEV). METHODS: A total of 396/596 patients (66%) from the prospective randomized phase II/III EORTC-26101 trial (with n=242 in the BEV and n=154 in the non-BEV arm) met the inclusion criteria with availability of anatomical and diffusion MRI-sequences at baseline prior treatment. Apparent diffusion coefficient (ADC) histograms from the contrast-enhancing tumor volume were fitted to a double Gaussian distribution and the mean of the lower curve (ADClow) was used for further analysis. The predictive ability of ADClow was assessed with biomarker threshold models and multivariable Cox-regression for overall and progression-free survival (OS, PFS). RESULTS: ADClow was associated with PFS (HR=0.625,p=0.007) and OS (HR=0.656,p=0.031). However, no (predictive) interaction between ADClow and the treatment arm was present (p=0.865 for PFS, p=0.722 for OS). Independent (prognostic) significance of ADClow was retained after adjusting for epidemiological, clinical and molecular characteristics (p 0.02 for OS, p 0.01 PFS). The biomarker threshold model revealed an optimal ADClow cutoff of 1241*10-6mm²/s for OS. Thereby, median OS for BEV-patients with ADClow 1241 was 10.39 months vs. 8.09 months for those with ADClow<1241 (p=0.004). Similarly, median OS for non-BEV patients with ADClow 1241 was 9.80 months vs. 7.79 months for those with ADClow<1241 (p=0.054). CONCLUSIONS: ADClow is an independent prognostic parameter for stratifying OS and PFS in patients with recurrent glioblastoma. Consequently, the previously suggested role of ADClow as predictive imaging biomarker could not be confirmed within this phase II/III trial.

DOI: https://doi.org/10.1093/neuonc/noaa120

Wick, Wolfgang; Bendszus, Martin; Kickingereder, Philipp (2020). Validation of diffusion MRI phenotypes for predicting response to bevacizumab in recurrent glioblastoma: post-hoc analysis of the EORTC-26101 trial. Neuro-Oncology, 22(11):1667-1676.
DOI: https://doi.org/10.1093/neuonc/noaa120

Validation of diffusion MRI phenotypes for predicting response to bevacizumab in recurrent glioblastoma: post-hoc analysis of the EORTC-26101 trial

Marianne Schell MD[1], Irada Pflüger MD[1], Gianluca Brugnara MD[1], Fabian Isensee MSc[2], Ulf Neuberger MD[1], Martha Foltyn MD[1], Tobias Kessler MD[3,4], Felix Sahm MD[5], Antje Wick MD[6], Martha Nowosielski MD PhD[3,6], Prof. Sabine Heiland PhD[1], Prof. Michael Weller MD[7], Prof. Michael Platten MD[8], Klaus H Maier-Hein PhD[2], Prof. Andreas von Deimling MD[5], Prof. Martin J. van den Bent MD[9], Thierry Gorlia PhD[10], Prof. Wolfgang Wick MD[3,4], Prof. Martin Bendszus MD[1], Philipp Kickingereder MD[1]

(1) Department of Neuroradiology, Heidelberg University Hospital, Heidelberg, Germany
(2) Medical Image Computing, German Cancer Research Center (DKFZ), Heidelberg, Germany
(3) Neurology Clinic, Heidelberg University Hospital, Heidelberg, Germany
(4) German Cancer Consortium (DKTK) in the German Cancer Research Center (DKFZ), Heidelberg, Germany
(5) Department of Neuropathology, Institute of Pathology, Heidelberg University Hospital, Heidelberg, Germany
(6) Department of Neurology, Medical University Innsbruck, Austria
(7) Department of Neurology, University Hospital and University of Zurich, Zurich, Switzerland
(8) Department of Neurology, Mannheim Medical Center, University of Heidelberg, Mannheim, Germany
(9) Brain Tumor Center at Erasmus MC Cancer Institute, Rotterdam, Netherlands
(10) European Organization for Research and Treatment of Cancer (EORTC), Brussels, Belgium

Corresponding Author and Address for Reprint Requests:

Philipp Kickingereder, MD MBA

Department of Neuroradiology, Heidelberg University Hospital

Im Neuenheimer Feld 400, 69120 Heidelberg, Germany

Email: philipp.kickingereder@med.uni-heidelberg.de

Phone: +49 (0) 6221 56 39069, Fax: +49 (0) 6221 56 4673

Abstract length: 290 words; **Manuscript length**: 3170 words; **Figures**: 6; **Tables**: 2; **References**: 35

Funding: Funded by the Else-Kröner Memorial Scholarship of the Else Kröner-Fresenius Foundation and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 404521405, SFB 1389 - UNITE Glioblastoma, Work Package C02 and Priority Programme 2177 “Radiomics: Next Generation of Biomedical Imaging”

Abstract

Background

In this study we validate a previously described diffusion-MRI phenotype as a potential predictive imaging biomarker for benefit from bevacizumab (BEV) in patients with recurrent glioblastoma.

Methods

MRI-data from the prospective randomized phase II/III EORTC-26101 trial in patients with first recurrence of glioblastoma were analyzed. A total of 396/596 patients (66%, with 242 in the BEV and 154 in the non-BEV treatment arm) met the inclusion criteria with availability of anatomical and diffusion MRI-sequences at baseline prior to treatment. The apparent diffusion coefficient (ADC) histograms from the contrast-enhancing tumor volume were fitted to a mixed model with a double Gaussian distribution and the mean of the lower curve (ADC_{low}) was used for further analysis. The predictive ability of ADC_{low} was assessed with biomarker threshold models and multivariable Cox-regression for OS and for progression-free survival (PFS).

Results

ADC_{low} was associated with PFS (HR=0.625 [95%CI 0.445-0.880], p=0.007) and OS (HR=0.656 [95%CI 0.448-0.962], p=0.031). However, no (predictive) interaction between ADC_{low} and the treatment arm was present (p=0.865 for PFS, p=0.722 for OS). Independent (prognostic) significance of ADC_{low} was retained after adjusting for epidemiological, clinical and molecular characteristics (p≤0.02 for OS and ≤0.01 PFS). The biomarker threshold model revealed an optimal ADC_{low} cutoff of $1241*10^{-6}$ mm²/s for OS. Thereby, median OS for BEV-patients with ADC_{low}<1241 was 10.39 months (95% CI, 9.4-15.55 months) vs. 8.09 months (95%CI, 7.1-9.53 months) for those with ADC_{low}≥1241 (p=0.004). Similarly, median OS for non-BEV patients with ADC_{low}<1241 was 9.80 months (95%CI, 8.94-15.20 months) vs. 7.79 months (95%CI, 6.15-9.50 months) for those with ADC_{low}≥1241 (p=0.054).

Conclusions

ADC_{low} is an independent prognostic parameter for stratifying OS and PFS in patients with recurrent glioblastoma. Consequently, the previously suggested role of ADC_{low} as predictive imaging biomarker for benefit from BEV could not be confirmed within this randomized controlled phase II/III trial.

Introduction

Glioblastoma remains one of the most aggressive forms of malignant primary brain tumor in adults [1]. Following standard treatment with maximal safe surgical resection, radiotherapy plus concomitant and maintenance temozolomide chemotherapy, almost all patients experience tumor progression. One commonly used treatment option at recurrence is bevacizumab (BEV), a humanized monoclonal antibody against vascular endothelial growth factor (VEGF) to block angiogenesis. BEV was approved for the treatment of recurrent glioblastoma by the US Food and Drug Administration (FDA) on the basis of two uncontrolled phase-II trials that demonstrated durable radiographic and, more importantly, clinical benefit in many patients [2, 3]. Although there is much support for the use of BEV, randomized phase-III trials (AVAglio, RTOG-0825, EORTC-26101) conducted to date have failed to show an overall survival (OS) benefit for BEV, both in the setting of newly-diagnosed [4, 5] and recurrent glioblastoma [6]. However, it seems to be reasonable to hypothesize that there may be subgroups of patient with a meaningful clinical therapy benefit and post-hoc analysis of the AVAglio trial suggested that the presence of a proneural molecular subtype is a predictive biomarker for stratifying BEV-treatment outcome in terms of OS [7]. However, there remains an unmet clinical need for easily, ideally non-invasively accessible, surrogate biomarkers able to delineate molecular activity and predict outcome to bevacizumab treatment [8-10]. Non-invasive characterization of glioblastoma on magnetic resonance imaging (MRI) has shown great potential for therapy guidance and previous studies indeed suggested that patients with a specific phenotype on diffusion MRI may derive a significant OS benefit from BEV treatment [11-15]. Specifically, histogram analysis of apparent diffusion coefficient values (ADC) from diffusion MRI with fitting of a double Gaussian mixed model and quantification of the mean ADC from the lower curve (ADC_{low}) has been suggested as a potential predictive imaging biomarker for stratifying BEV-treatment outcome in terms of OS, both in the setting of newly diagnosed [14, 15] and recurrent glioblastoma [11-13]. However, despite these intriguing findings the utility of ADC_{low} as a truly predictive imaging biomarker has not yet been fully established, since available studies are based on the analysis of retrospective datasets [12-

15] or post-hoc analysis of available phase II trials [11]. Specifically, data from randomized phase III trials with availability of a BEV-free control arm would be necessary to reliable differentiate whether ADC_{low} is a truly predictive biomarker (and allows to identify the likelihood of sensitivity to BEV treatment) or whether it only serves as a prognostic biomarker (and thus reflects the patients overall outcome, regardless of therapy) [16].

Here we performed an exploratory post-hoc analysis of the randomized controlled multicentre phase II and III EORTC-26101 trial (NCT01290939) in patients with first recurrence of glioblastoma. Specifically, we assessed whether ADC_{low} quantified from baseline MRI prior to treatment is a prognostic and/or predictive imaging biomarker for stratifying treatment response to BEV (in terms of OS and PFS).

Methods

Study design and participants

For the present study we analyzed MRI data from brain tumor patients acquired as part of the European Organisation for Research and Treatment of Cancer (EORTC) 26101 trial (NCT01290939). As described previously [17], the EORTC-26101 study was a prospective randomized phase II and III trial in patients with first progression of a glioblastoma after standard chemo-radiotherapy. In brief, the phase II trial evaluated the optimal treatment sequence of bevacizumab and lomustine (four treatment arms with single agent vs. sequential vs. combination) whereas the subsequent phase III trial (two treatment arms) compared patients treated with lomustine alone with those receiving a combination of lomustine and bevacizumab. Overall, the EORTC-26101 study included 596 patients (n=159 from phase II and n=437 from phase III) at 37 institutions within Europe. The study was conducted in accordance with the Declaration of Helsinki and the protocol was approved by local ethics committees and patients provided written informed consent (EudraCT# 2010-023218-30). Full study design and outcomes have been published previously [7, 16]. MRI exams were acquired with a standardized imaging protocol with acquisition of pre- and postcontrast T1-weighted, T2-weighted and FLAIR images at baseline and every 6 weeks until week 24, afterwards every 3 months according to a prespecified MR protocol [18]. In addition, optional acquisition of diffusion-weighted MRI (DWI) was performed at the discretion of the individual institution.

Based on the four treatment arms of the phase II trial and the two treatment arms in the phase III trial patients were allocated for the present post-hoc analysis into two groups (a) those receiving bevacizumab (BEV subgroup) and, (b) those not receiving bevacizumab (non-BEV subgroup) for the treatment at first recurrence of glioblastoma. Specifically, the BEV subgroup consisted of patients that were allocated to the bevacizumab plus lomustine treatment arm (at progression salvage treatment at the investigators best choice) or the bevacizumab alone treatment arm (at progression switch to bevacizumab plus lomustine), whereas the non-BEV subgroup consisted of patients that were allocated to the lomustine alone treatment arms (at

progression either switch to bevacizumab or salvage treatment at the investigators best choice).

Sample size for the present analysis was determined by the availability of MRI data and not derived from a power calculation. Specifically, we required MRI examinations that included (a) anatomical sequences (pre- and postcontrast T1-weighted [T1-w, cT1-w], T2-weighted [T2-w] and FLAIR images) and (b) DWI or ADC sequences at baseline prior treatment. We excluded those MRI examinations with (a) corrupt data following DICOM to *NIfTI* conversion (primarily due to the non-standardized center-specific anonymization of DICOM-files), (b) incomplete availability of T1-w, cT1-w, FLAIR, T2-w and DWI or ADC sequences (requiring either 3D acquisitions or 2D with axial orientation), (c) heavy motion artifacts, or (d) insufficient contrast-agent administration during cT1-w acquisition.

Procedures

The ADC sequences were supplied by the individual institution (calculated on-the-fly by the scanner software). For those cases with missing ADC but available DWI, we generated ADC maps from DWI using dedicated software (Olea Sphere v 2.3, Olea Medical, La Ciotat, France). Next, processing of the anatomical sequences (T1-w, cT1-w, T2-w and FLAIR) was performed as described previously [19, 20] and included automated deep-learning based brain extraction [20] followed by registration of the brain-extracted cT1-w, FLAIR, T2-w image volumes to the respective brain extracted T1-w image volume using the linear image registration tool (FMRIB software library, FSL, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSL) with spline interpolation and a 6-degree of freedom transformation. Subsequently automated deep-learning based voxel-wise segmentation and volumetric quantification of the contrast-enhancing tumor volume was performed [19]. Next, the binary brain mask from the T1-w image volume was transferred to the ADC image volumes, followed by registration of the brain-extracted ADC image volumes to the respective brain extracted T1-w image volume using the linear image registration tool (FMRIB software library, FSL, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSL) with spline interpolation and a 6-degree of freedom transformation. The transformation matrix that was generated

during registration of the brain-extracted ADC to the brain-extracted T1-w image volume was inverted and applied to the tumor segmentation mask. Thereby, the tumor segmentation mask was transformed to the original (non-registered) image space of the ADC sequence using nearest neighbor interpolation.

Subsequent processing was done using MATLAB R2018a (The MathWorks, Inc., Natick, Massachusetts, United States). First, extraction of ADC values within the contrast-enhancing tumor segmentation mask of the original (non-registered) ADC sequence was performed and all extracted ADC values were used to construct the histogram. An automatic binning algorithm was used to determine the optimal number and width of the bins. A two-term Gaussian model with a robust non-linear least square fitting method and a trust-region algorithm was applied by using the Curve Fitting Toolbox (ver. 3.5.5). We defined the following fitting options: (1) constraining the lower bounds on the coefficients, specifically the peaks to 1 voxel and the means of the Gaussian curves to the lowest voxel value in the contrast-enhancing tumor segmentation mask, (2) constraining the upper bound on the coefficients, specifically the peaks to the number of voxel within the mask and the means to highest voxel value in the contrast-enhancing tumor segmentation mask, (3) determine the starting points of the lower centroid at the 25^{th} and upper centroid at the 75^{th} percentile. The goodness of fit was evaluated with R^2 and through visual inspection of the residuals (cases with R^2 values <0.7 were excluded from further analysis). No further manual parameter adjustments were applied. The mean of the lower Gaussian curve (ADC_{low}) was used for subsequent statistical analysis. Figure 1 illustrates the workflow in a representative case.

Outcomes

The objective of the present study was to identify whether ADC_{low} is a prognostic and/or predictive imaging biomarker for stratifying treatment outcome (by means of PFS and OS) to BEV therapy in patients with recurrent glioblastoma. OS was calculated from the date of randomization until death or last follow-up. PFS was calculated from the date of randomization until the date of tumor progression (as assessed by local RANO reading as part of the initial analysis of the EORTC-26101 trial [6]) or death (whichever occurred first).

Statistical analysis

Statistical analysis was performed with R version 3.6.2 (R Foundation for Statistical Computing, Vienna, Austria). The distribution of baseline clinical, molecular and imaging characteristics between the BEV and non-BEV groups was compared with the chi-square test for categorical parameters and the Wilcoxon rank sum test for continuous parameters.

The potential predictive association between ADC_{low} and the effect of BEV was evaluated by constructing predictive biomarker threshold models. ADC_{low} was fitted using a hierarchical Bayesian method under the framework of the Cox proportional hazards model separately for OS and PFS (implemented with the bhm package in R [https://cran.r-project.org/web/packages/bhm] using standard parameters i.e. with B=500 burn-in steps and R=2000 replications for the Bayesian method as recommended by the developers for the analysis of a biomarker variable) [21]. The hierarchical Bayesian method simultaneously makes a statistical inference on the ADC_{low} threshold defining the sensitive patient subset and the interaction effect in terms of outcome (OS or PFS) between ADC_{low} and the treatment (i.e. BEV vs. non-BEV subgroup).

Multivariable Cox proportional hazards regression modeling with the continuous ADC_{low} value for both OS and PFS was separately performed for the BEV and non-BEV groups. The model was adjusted for potential clinical confounders (age, sex, WHO performance status, glucocorticoid administration at randomization, sequence of treatment), molecular confounders (O^6-methylguanine-DNA methyltransferase (MGMT) promoter methylation status, glioma-CpG island methylator phenotype (GCIMP) status)) and baseline contrast-enhancing tumor volume. The Kaplan-Meier method with the log-rank test was used to construct survival curves for OS and PFS in both BEV and non-BEV subgroups (and grouping patients based on the cut-off determined by the biomarker threshold model for ADC_{low} value). The association between ADC_{low} values and MGMT promotor methylation status was examined with the Wilcoxon rank sum test. P-values <0.05 were considered significant.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. MS, MB and PK had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Out of 596 patients in the EORTC-26101 phase II and III study, a subset of 396 patients (66.4%) met the outlined criteria for this post-hoc analysis (see Figure 2). Specifically, the 396 patients included 242 patients (61.1%) in the BEV subgroup and 154 patients (38.9%) in the non-BEV subgroup. Baseline clinical characteristics (sex, age, WHO performance status, use of glucocorticoids, sequence of treatment), molecular characteristics (MGMT promoter methylation status, GCIMP status) and contrast-enhancing tumor volumes were equally distributed between the BEV and non-BEV subgroup (Table 1). The ADC_{low} values were balanced with a median of $1027*10^{-6}$ mm²/s for the BEV subgroup and $1125*10^{-6}$ mm²/s for the non-BEV subgroup (p=0.46, see Table 1). Moreover, R_2 (as a measure of goodness-of-fit) showed no difference between the groups, resulting in an unbiased fitting result for the comparison (p=0.998) (see Supplementary Table S1).

PFS was longer in the BEV group (3.98 months [95% CI, 3.22-4.21 months]) than in the non-BEV group (1.55 months [95% CI, 1.48-2.60 months], p<0.0001), whereas OS was not different between the BEV group (8.91 months [95% CI, 7.99-9.96 months]) and the non-BEV group (8.61 months [95% CI, 7.76-9.80 months], p=0.95) (see Figure 3). The biomarker threshold models demonstrated that ADC_{low} was significantly associated with both PFS (HR=0.625 [95% CI 0.445-0.880], p=0.007) and OS (HR=0.656 [95% CI 0.448-0.962], p=0.031). However, no (predictive) interaction between ADC_{low} and the treatment arm was present, neither for PFS (HR=1.04 [95% CI 0.66-1.63], p=0.865) nor for OS (HR=0.91 [95% CI 0.54-1.54], p=0.7221) (see Table 2). Multivariable Cox-regression modeling confirmed the independent (prognostic) significance of ADC_{low} after adjusting for epidemiological, clinical and molecular characteristics (p≤0.02 for OS and ≤0.01 PFS) (see Figure 4).

The biomarker threshold model revealed an optimal ADC_{low} cutoff of $1241*10^{-6}$ mm²/s for OS (corresponding to a 75:25 split of the dataset, see Table 2). By applying this cutoff, the median OS for BEV-patients with ADC_{low}<1241 was 10.39 months (95% CI, 9.4-15.55 months) vs. 8.09 months (95% CI, 7.1-9.53 months) for those with ADC_{low}>1241 (p=0.004) (see Figure 5).

Kommentiert [WM1]: Reicht nicht eine Stelle hinter dem Komma?

Kommentiert [WM2]: 2 stellen hinter dem Komma sind genug

Similarly, median OS for non-BEV patients with ADC_{low}<1241 was 9.80 months (95%CI, 8.94-15.20 months) vs. 7.79 months (95%CI, 6.15-9.50 months) for those with ADC_{low}>1241 (p=0.054) (see Figure 5). For PFS the biomarker threshold model revealed an optimal ADC_{low} cutoff of $1217*10^{-6}$ mm^2/s (corresponding to a 70:30 split of the dataset, see Table 2). By applying this cutoff, the median PFS for BEV-patients with ADC_{low}<1241 was 5.38 months (95% CI, 4.21-6.58 months) vs. 3.29 months (95% CI, 2.89 – 4.14 months) for those with ADC_{low}>1241 (p=0.003) (see Figure 5). Similarly, median PFS for non-BEV patients with ADC_{low}<1241 was 2.73 months (95% CI, 1.55-4.21 months) vs. 1.48 months (95% CI, 1.45-1.64 months) for those with ADC_{low}>1241 (p=0.021) (see Figure 5).

There was no association between ADC_{low} and MGMT promoter methylation status (p=0.13) (see Figure 6). Specifically, median ADC_{low} values were $1124*10^{-6}$ mm^2/s (IQR, $985*10^{-6}$ mm^2/s - $1280*10^{-6}$ mm^2/s) for MGMT promoter-methylated and $1085*10^{-6}$ mm^2/s (IQR, $951*10^{-6}$ mm^2/s - $1233*10^{-6}$ mm^2/s) for MGMT promoter-unmethylated tumors.

Kommentiert [WM3]: MGMT could also go to the Supplement….. no strong biological hypothesis and negative data anyway

Discussion

Non-invasive characterization of glioblastoma on MRI has been extensively used for studying the effects of BEV and for attempting non-invasive identification of patient subsets for whom anti-angiogenic treatment may be beneficial [11, 22, 23]. In this context, the discovery of ADC_{low} as a potential predictive imaging biomarker for stratifying both PFS and OS has gained significant interest [11-15]. Specifically, initial studies obtained from retrospective datasets demonstrated that both newly diagnosed as well as recurrent glioblastoma patients with lower ADC_{low} levels exhibit a more durable response from BEV, both in terms of PFS and OS [12-15]. Subsequent post-hoc analysis of the randomized controlled phase II BELOB trial confirmed these earlier findings [11]. In particular, the BELOB trial compared the use of single-agent bevacizumab vs. lomustine vs. a combination of bevacizumab plus lomustine [24]. ADC_{low} was significantly associated with OS in patients treated with single-agent BEV but not in those with single-agent lomstine; no results were shown for the combination arm of bevacizumab plus lomustine [11]. Based on these results it was suggested that ADC_{low} is a predictive imaging biomarker and allows to identify the likelihood of sensitivity to BEV treatment [11]. However, despite these intriguing findings, confirmatory evidence from phase III trials is lacking [4-6]. In the present post-hoc analysis of the randomized controlled multicentre phase II and III EORTC-26101 trial we therefore aimed to validate the utility of ADC_{low} as a predictive imaging biomarker. By applying a robust post-processing workflow, we found that ADC_{low} was significantly associated with PFS and OS, however no (predictive) interaction between ADC_{low} and treatment group (BEV vs. non-BEV) was present. Consequently, these findings challenge previous studies and suggest that ADC_{low} serves as a prognostic biomarker and thus reflects the patient's overall outcome, regardless of therapy in the setting of recurrent glioblastoma. One could argue that the lack of predictive association between ADC_{low} and BEV-treatment in terms of OS within the scope of the EORTC-26101 trial was confounded by cross-over between treatment arms. Specifically, cross-over was seen in 35% of patients within the non-BEV subgroup who received BEV treatment following disease progression. Nevertheless, this effect may have been negligible given the identical (prognostic) pattern that was found for

ADC_{low} also in terms of PFS (which was not affected by any cross-over). Moreover, the sequence of treatment had no independent significance on multivariable Cox regression modeling. The cut-offs for ADC_{low} that were suggested by the biomarker threshold model (OS: $1241*10^{-6}$ mm²/s and PFS: $1217*10^{-6}$ mm²/s) were comparable to previous works that obtained cutoffs of $1200*10^{-6}$ mm²/s [12-15] and $1240*10^{-6}$ mm²/s [11], implying a robust extraction independent of scanner hardware and extraction software. A possible explanation for the prognostic effect of ADC_{low} would be the well-known association of tumor cellularity and ADC [25, 26]. Lower ADC values are associated with a higher cellularity and higher tumor burden [27-29], which could explain a prognostic effect of ADC_{low} independent of treatment [30-32]. In contrast, mechanism of ADC_{low} as a predictive biomarker for BEV remain speculative and unclear [11].

Given the prognostic importance of ADC_{low} and MGMT promoter methylation status, we specifically investigated their relationship, since previous studies suggested lower ADC_{low} values for newly-diagnosed MGMT-methylated glioblastoma Pope, Lai [15] whereas other studies showed no association between ADC measurements and MGMT status [33-35]. Our study also did not show any association between ADC_{low} and MGMT promoter methylation status, thereby suggesting that ADC_{low} and MGMT status contribute to prognosis of recurrent glioblastoma through independent mechanisms with additive prognostic effects.

Our study has some limitations. First, we acknowledge that only a subset of 396 of 596 patients (66%) met the criteria for this post-hoc analysis with sample size determined by the availability of MRI data and not derived from a power calculation. Second, we could not perform central standardized calculation of ADC sequences for all analyzed patients, but instead relied on the ADC sequences that were generated on-the-fly by the scanner software of the individual institutions for most cases. This may have led to a non-standardized generation of ADC sequences due to different scanner software. However, the applied ADC histogram analysis has proven to be a very robust method with high reproducibility [36]. Third, the information on b-values used for acquiring the DWI sequences was missing for most cases (due to routine

Kommentiert [WM4]: Findet man das in der „Results“ section?

Kommentiert [WM5]: Formatting Pope Lai

Kommentiert [WM6]: Man könnte in einer S-Table zeigen, dass sich die beiden Gruppen der 396 mit Bildern nicht prognostisch/klinisch von den beiden Gruppen der 200 ohne Bilder unterscheiden

anonymization of relevant DICOM tags in multicenter studies). For those cases with available information on B-values we noted that acquisition was either performed with b=1000 or b=1200. Although it is well known that increasing b-values shift the ADC histograms to lower values it remains unlikely that this was a relevant confounder for not identifying a predictive, but only a prognostic association of ADC_{low} within the EORTC-26101 trial.

In conclusion, ADC_{low} is an independent prognostic parameter for stratifying OS and PFS within the scope of the EORTC-26101 trial in patients with recurrent glioblastoma. Consequently, the previously suggested role of ADC_{low} as predictive imaging biomarker was not confirmed. Overall, identifying patient subsets that may benefit from BEV remains challenging and the optimal indication for BEV in the treatment of glioblastoma remains uncertain.

References

1. Stupp, R., et al., *Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma.* N Engl J Med, 2005. **352**(10): p. 987-96.
2. Friedman, H.S., et al., *Bevacizumab alone and in combination with irinotecan in recurrent glioblastoma.* J Clin Oncol, 2009. **27**(28): p. 4733-40.
3. Kreisl, T.N., et al., *Phase II trial of single-agent bevacizumab followed by bevacizumab plus irinotecan at tumor progression in recurrent glioblastoma.* J Clin Oncol, 2009. **27**(5): p. 740-5.
4. Chinot, O.L., et al., *Bevacizumab plus radiotherapy-temozolomide for newly diagnosed glioblastoma.* N Engl J Med, 2014. **370**(8): p. 709-22.
5. Gilbert, M.R., et al., *A randomized trial of bevacizumab for newly diagnosed glioblastoma.* N Engl J Med, 2014. **370**(8): p. 699-708.
6. Wick, W., et al., *Lomustine and Bevacizumab in Progressive Glioblastoma.* N Engl J Med, 2017. **377**(20): p. 1954-1963.
7. Sandmann, T., et al., *Patients With Proneural Glioblastoma May Derive Overall Survival Benefit From the Addition of Bevacizumab to First-Line Radiotherapy and Temozolomide: Retrospective Analysis of the AVAglio Trial.* J Clin Oncol, 2015. **33**(25): p. 2735-44.
8. Nowosielski, M., et al., *Progression types after antiangiogenic therapy are related to outcome in recurrent glioblastoma.* Neurology, 2014. **82**(19): p. 1684-92.
9. Kickingereder, P., et al., *Relative cerebral blood volume is a potential predictive imaging biomarker of bevacizumab efficacy in recurrent glioblastoma.* Neuro Oncol, 2015.
10. Kickingereder, P., et al., *MR Perfusion-derived Hemodynamic Parametric Response Mapping of Bevacizumab Efficacy in Recurrent Glioblastoma.* Radiology, 2015: p. 151172.

11. Ellingson, B.M., et al., *Diffusion MRI Phenotypes Predict Overall Survival Benefit from Anti-VEGF Monotherapy in Recurrent Glioblastoma: Converging Evidence from Phase II Trials.* Clin Cancer Res, 2017. **23**(19): p. 5745-5756.
12. Ellingson, B.M., et al., *Pretreatment ADC Histogram Analysis Is a Predictive Imaging Biomarker for Bevacizumab Treatment but Not Chemotherapy in Recurrent Glioblastoma.* American Journal of Neuroradiology 2014. **35**(4): p. 673-679.
13. Pope, W.B., et al., *Apparent diffusion coefficient histogram analysis stratifies progression-free and overall survival in patients with recurrent GBM treated with bevacizumab: a multi-center study.* J Neurooncol, 2012. **108**(3): p. 491-8.
14. Pope, W.B., et al., *Recurrent Glioblastoma Multiforme: ADC Histogram Analysis Predicts Response to Bevacizumab Treatment.* Radiology, 2009. **252**(1): p. 182-189.
15. Pope, W.B., et al., *Apparent diffusion coefficient histogram analysis stratifies progression-free survival in newly diagnosed bevacizumab-treated glioblastoma.* AJNR Am J Neuroradiol, 2011. **32**(5): p. 882-9.
16. Ballman, K.V., *Biomarker: Predictive or Prognostic?* Journal of Clinical Oncology, 2015. **33**(33): p. 3968-3971.
17. Kickingereder, P., et al., *Automated quantitative tumor response assessment of MRI in neuro-oncology with artificial neural networks: a multicenter, retrospective study.* Lancet Oncol, 2019 (in press).
18. Ellingson, B.M., et al., *Consensus recommendations for a standardized Brain Tumor Imaging Protocol in clinical trials.* Neuro Oncol, 2015. **17**(9): p. 1188-98.
19. Kickingereder, P., et al., *Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study.* Lancet Oncol, 2019. **20**(5): p. 728-740.
20. Isensee, F., et al. *Automated brain extraction of multi-sequence MRI using artificial neural networks*. arXiv e-prints, 2019.

21. Chen, B.E., W. Jiang, and D. Tu, *A hierarchical Bayes model for biomarker subset effects in clinical trials.* Computational Statistics & Data Analysis, 2014. **71**: p. 324-334.
22. Mayer, T.M., *Can We Predict Bevacizumab Responders in Patients With Glioblastoma?* J Clin Oncol, 2015. **33**(25): p. 2721-2.
23. Kickingereder, P., et al., *Large-scale Radiomic Profiling of Recurrent Glioblastoma Identifies an Imaging Predictor for Stratifying Anti-Angiogenic Treatment Response.* Clin Cancer Res, 2016. **22**(23): p. 5765-5771.
24. Taal, W., et al., *Single-agent bevacizumab or lomustine versus a combination of bevacizumab plus lomustine in patients with recurrent glioblastoma (BELOB trial): a randomised controlled phase 2 trial.* Lancet Oncol, 2014. **15**(9): p. 943-53.
25. Eidel, O., et al., *Automatic Analysis of Cellularity in Glioblastoma and Correlation with ADC Using Trajectory Analysis and Automatic Nuclei Counting.* PLoS One, 2016. **11**(7): p. e0160250.
26. Surov, A., H.J. Meyer, and A. Wienke, *Correlation between apparent diffusion coefficient (ADC) and cellularity is different in several tumors: a meta-analysis.* Oncotarget, 2017. **8**(35): p. 59492-59499.
27. Gauvain, K.M., et al., *Evaluating pediatric brain tumor cellularity with diffusion-tensor imaging.* AJR Am J Roentgenol, 2001. **177**(2): p. 449-54.
28. Gupta, R.K., et al., *Relationships between choline magnetic resonance spectroscopy, apparent diffusion coefficient and quantitative histopathology in human glioma.* J Neurooncol, 2000. **50**(3): p. 215-26.
29. Sugahara, T., et al., *Usefulness of diffusion-weighted MRI with echo-planar technique in the evaluation of cellularity in gliomas.* J Magn Reson Imaging, 1999. **9**(1): p. 53-60.
30. Higano, S., et al., *Malignant astrocytic tumors: clinical importance of apparent diffusion coefficient in prediction of grade and prognosis.* Radiology, 2006. **241**(3): p. 839-46.

31. Murakami, R., et al., *Malignant supratentorial astrocytoma treated with postoperative radiation therapy: prognostic value of pretreatment quantitative diffusion-weighted MR imaging.* Radiology, 2007. **243**(2): p. 493-9.
32. Zulfiqar, M., D.M. Yousem, and H. Lai, *ADC values and prognosis of malignant astrocytomas: does lower ADC predict a worse prognosis independent of grade of tumor?--a meta-analysis.* AJR Am J Roentgenol, 2013. **200**(3): p. 624-9.
33. Choi, Y.S., et al., *Incremental Prognostic Value of ADC Histogram Analysis over MGMT Promoter Methylation Status in Patients with Glioblastoma.* Radiology, 2016. **281**(1): p. 175-84.
34. Moon, W.J., et al., *Imaging parameters of high grade gliomas in relation to the MGMT promoter methylation status: the CT, diffusion tensor imaging, and perfusion MR imaging.* Neuroradiology, 2012. **54**(6): p. 555-63.
35. Gupta, A., et al., *Diffusion-weighted MR imaging and MGMT methylation status in glioblastoma: a reappraisal of the role of preoperative quantitative ADC measurements.* AJNR Am J Neuroradiol, 2013. **34**(1): p. E10-1.
36. Steens, S.C., et al., *Reproducibility of brain ADC histograms.* Eur Radiol, 2004. **14**(3): p. 425-30.

Tables

Table 1. Baseline characteristics of patients included in the present study.

Parameter	BEV group (n=242)	non-BEV group (n=154)	p-value
Clinical characteristics			
Sex — no. (%)			
Female	95 (39)	54 (35)	0.46
Male	147 (61)	400 (65)	
Age — yr.			
Median	58	60	0.24
Interquartile range	51, 65	52, 66	
WHO performance status — no. (%)			
0	79 (33)	51 (33)	1.00
≥1	163 (67)	101 (67)	
Use of glucocorticoids — no. (%)			
No	112 (46)	79 (51)	0.38
Yes	130 (54)	75 (49)	
Sequence of treatment — no. (%)			
BEV (at PD: BEV+LOM)	54 (22)	-	n.a.
BEV+LOM (at PD investigators choice)	188 (78)	-	
LOM (at PD: BEV)	-	54 (35)	
LOM (at PD investigators choice)	-	100 (65)	
Molecular characteristics			
MGMT promoter methylation status — no. (%)			
Methylated	74 (31)	54 (35)	0.65
Unmethylated	77 (32)	46 (30)	
Undetermined / Missing data	91 (38)	54 (35)	
GCIMP status — no. (%)			
Positive	5 (2)	7 (5)	0.27
Negative	237 (98)	147 (95)	
Quantitative MRI parameters at baseline MRI			
Contrast-enhancing tumor volume (cm^3)			
Median	15.38	17.93	0.25
Interquartile range	6.80, 29.58	8.70, 32.43	
ADC_{low} (10^{-6} mm^2/s)			
Median	1078	1125	0.46
Interquartile range	954, 1229	987, 1272	

Abbreviations: ADC=apparent diffusion coefficient, BEV=bevacizumab, LOM=lomustine, PD=progressive disease, G-CIMP = glioma-CpG island methylator phenotype, MGMT = O6-methylguanine-DNA methyltransferase, n.a. = not assessed

Table 2: Results of biomarker threshold model for overall survival (OS) and progression free survival (PFS). The cutoff refers to the optimal threshold suggest by the model. HR=hazard ratio.

	Parameter	HR	(95% CI)	P-value	Cut-off
OS	ADC_{low}	0.656	(0.448-0.962)	0.031	$1241*10^{-6}$ mm^2/s
	Treatment	0.984	(0.757-1.280)	0.905	
	ADC_{low}:Treatment*	0.909	(0.537-1.538)	0.722	
PFS	ADC_{low}	0.625	(0.445-0.880)	0.007	$1217*10^{-6}$ mm^2/s
	Treatment	0.517	(0.402-0.666)	<0.001	
	ADC_{low}:treatment*	1.040	(0.662-1.634)	0.865	

Annotation: * interaction term

Figures

Figure 1. Processing pipeline for a representative patient. (A) original MR images (axial slice of contrast-enhanced T1-weighted, T1-weighted, T2-weighted and FLAIR image). (B) brain-extracted, contrast-enhanced T1-weighted image with mask of contrast-enhanced lesion (light blue). (C) ventral and lateral view of surface rendering with 3D contrast-enhanced lesion mask. (D) non-registered ADC map with applied contrast-enhanced lesion mask (light blue). (E) extracted histogram of ADC values with individual Gaussian curves (white broken line) and the double Gaussian fit (dark blue). Horizontal line with individual ADC_{low} value.

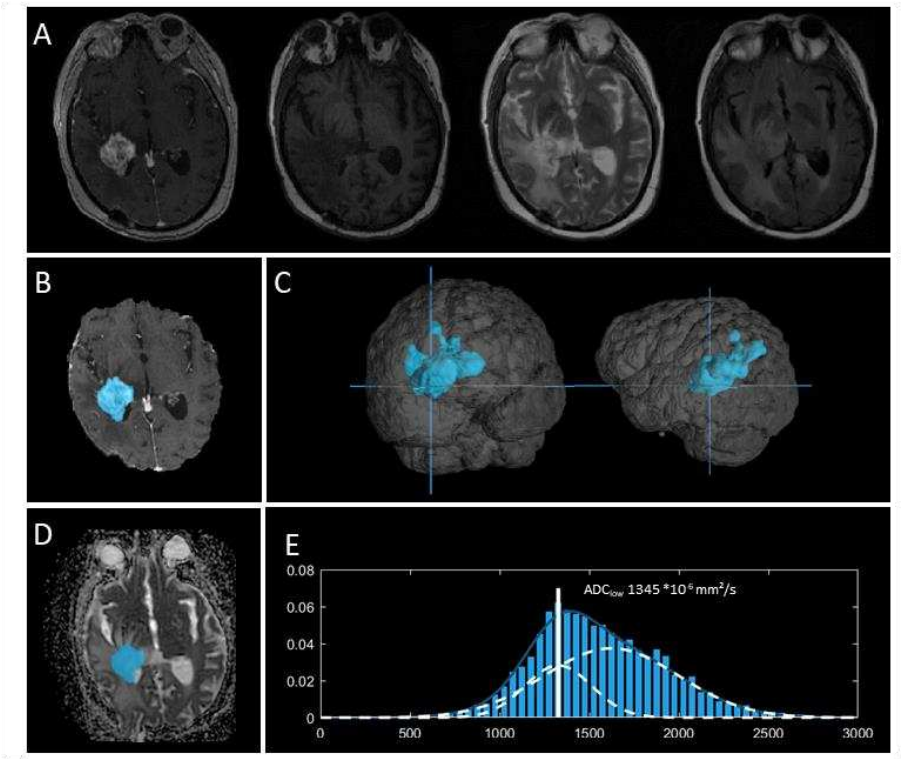


Figure 2. Flowchart of the procedures and analysis performed within the present study.

Kommentiert [WM7]: I would say BEV group and non-BEV group throughout because pts are pooled from different study arms which may cause confusion

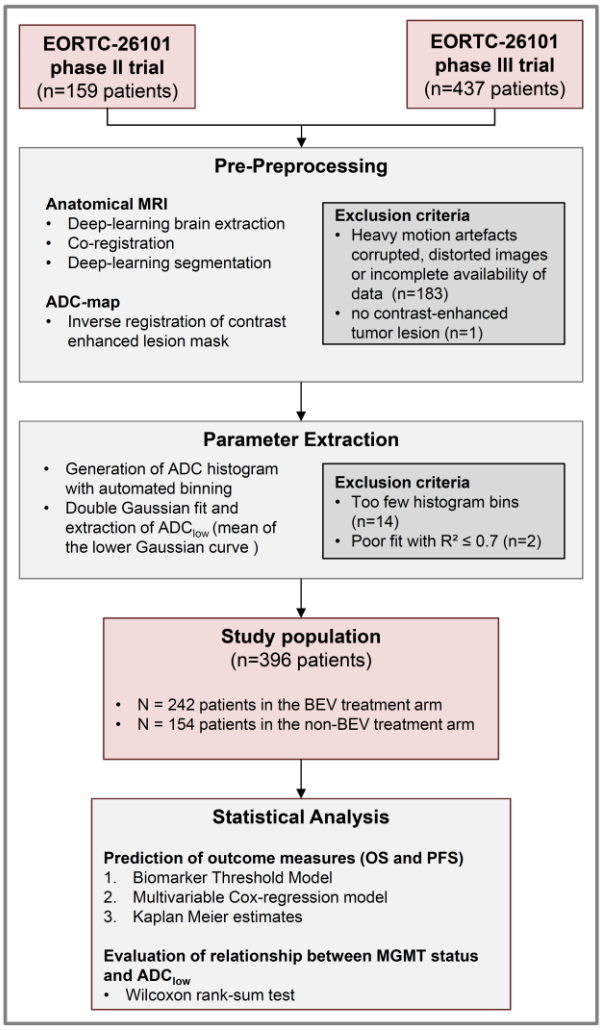
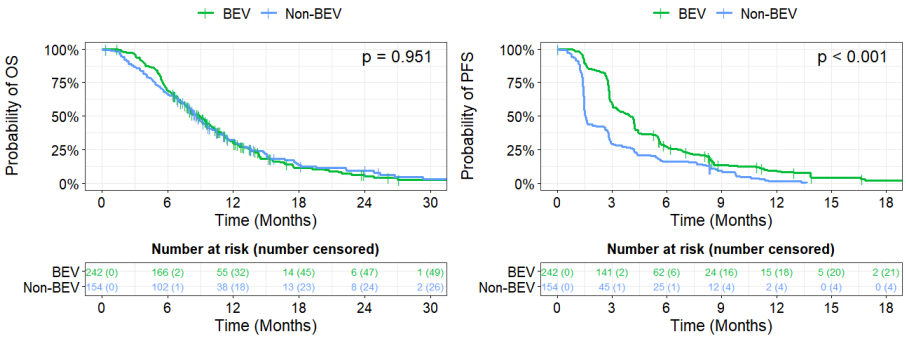


Figure 3. Kaplan-Meier plots for overall survival (OS, left panel) and progression free-survival (PFS, right panel) comparing BEV vs. non-BEV groups. Treatment with BEV was associated with longer PFS (p<0.001) but not OS (p=0.951).



Kommentiert [WM8]: Redundant if you provide p value

Kommentiert [WM9]: Blue/red is eeasier to distinguish

Figure 4. Multivariable Cox proportional hazards regression model for overall survival (OS, left panel) and progression free-survival (PFS, right panel) in BEV treated patients (upper panel) and non-BEV treated patients (lower panel). Independent significance of ADC_{low} was retained after adjusting for epidemiological, clinical and molecular characteristics.

Overall survival (OS): BEV subgroup

Variable	N	Hazard ratio		p
Baseline ADC-low	151		1.00 (1.00, 1.00)	0.017
Baseline CE tumor volume	151		1.02 (1.01, 1.03)	<0.001
Age	151		1.01 (0.99, 1.03)	0.423
Sex				
female	63		Reference	
male	88		1.58 (1.07, 2.35)	0.023
Corticosteroid intake				
no	71		Reference	
yes	80		1.21 (0.80, 1.83)	0.358
WHO Performance Status				
> 0	101		Reference	
0	50		0.89 (0.58, 1.37)	0.595
MGMT status				
Methylated	74		Reference	
Unmethylated	77		1.86 (1.25, 2.78)	0.002
GCIMP status				
negative	148		Reference	
positive	3		1.66 (0.21, 13.10)	0.630
Sequence of treatment				
BEV (at PD: BEV+LOM)	30		Reference	
BEV+LOM (at PD: inv. choice)	121		1.04 (0.65, 1.67)	0.871

0.5 1 2 5 10

Progression-free survival (PFS): BEV subgroup

Variable	N	Hazard ratio		p
Baseline ADC-low	151		1.00 (1.00, 1.00)	0.002
Baseline CE tumor volume	151		1.01 (1.00, 1.02)	0.058
Age	151		1.00 (0.98, 1.02)	0.805
Sex				
female	63		Reference	
male	88		1.44 (1.00, 2.06)	0.048
Corticosteroid intake				
no	71		Reference	
yes	80		1.29 (0.88, 1.89)	0.193
WHO Performance Status				
> 0	101		Reference	
0	50		1.14 (0.78, 1.67)	0.503
MGMT status				
Methylated	74		Reference	
Unmethylated	77		2.13 (1.47, 3.10)	<0.001
GCIMP status				
negative	148		Reference	
positive	3		0.52 (0.07, 3.99)	0.530
Sequence of treatment				
BEV (at PD: BEV+LOM)	30		Reference	
BEV+LOM (at PD: inv. choice)	121		0.85 (0.54, 1.34)	0.481

0.1 0.2 0.5 1 2

Overall survival (OS): non-BEV subgroup

Variable	N	Hazard ratio		p
Baseline ADC-low	100		1.00 (1.00, 1.00)	0.005
Baseline CE tumor volume	100		1.01 (1.00, 1.02)	0.011
Age	100		1.02 (1.00, 1.05)	0.088
Sex				
female	41		Reference	
male	59		0.84 (0.53, 1.35)	0.479
Corticosteroid intake				
no	52		Reference	
yes	48		2.04 (1.17, 3.55)	0.012
WHO Performance Status				
> 0	65		Reference	
0	35		0.87 (0.49, 1.54)	0.637
MGMT status				
Methylated	54		Reference	
Unmethylated	46		1.68 (1.04, 2.70)	0.033
GCIMP status				
negative	95		Reference	
positive	5		0.46 (0.15, 1.40)	0.172
Sequence of treatment				
LOM (at PD: BEV)	32		Reference	
LOM (at PD: inv. choice)	68		1.05 (0.64, 1.70)	0.854

0.2 0.5 1 2

Progression-free survival (PFS): non-BEV subgroup

Variable	N	Hazard ratio		p
Baseline ADC-low	100		1.00 (1.00, 1.00)	0.01
Baseline CE tumor volume	100		1.01 (1.00, 1.02)	0.29
Age	100		1.01 (0.99, 1.03)	0.54
Sex				
female	41		Reference	
male	59		0.71 (0.46, 1.09)	0.12
Corticosteroid intake				
no	52		Reference	
yes	48		1.10 (0.63, 1.92)	0.75
WHO Performance Status				
> 0	65		Reference	
0	35		0.87 (0.51, 1.51)	0.62
MGMT status				
Methylated	54		Reference	
Unmethylated	46		2.97 (1.84, 4.79)	<0.001
GCIMP status				
negative	95		Reference	
positive	5		1.92 (0.71, 5.20)	0.20
Sequence of treatment				
LOM (at PD: BEV)	32		Reference	
LOM (at PD: inv. choice)	68		1.01 (0.64, 1.60)	0.95

0.5 1 2 5

Kommentiert [WM10]: Die Zahlen unten für die HR sind zu klein

Figure 5. Kaplan-Meier plots for overall survival (OS, 1st row) and progression free-survival (PFS, 2nd row) stratifying based on their ADC_{low} value on baseline MRI. Patients with ADC_{low} values above the cut-off identified by the biomarker threshold model (≥1241*10^{-6} mm²/s for OS and ≥1217*10^{-6} mm²/s for PFS) demonstrated significantly longer PFS and OS as compared to patients with ADC_{low} values below the cut-off. This effect was seen for both BEV-subgroup (left column) and non-BEV subgroup (right column), thereby confirming that ADC_{low} is a prognostic and not a predictive imaging biomarker in the EORTC-26101 trial.

Kommentiert [WM11]: Is this not redundant with the text? It would suffice to state in the legend what is shown…. Results belong in the text

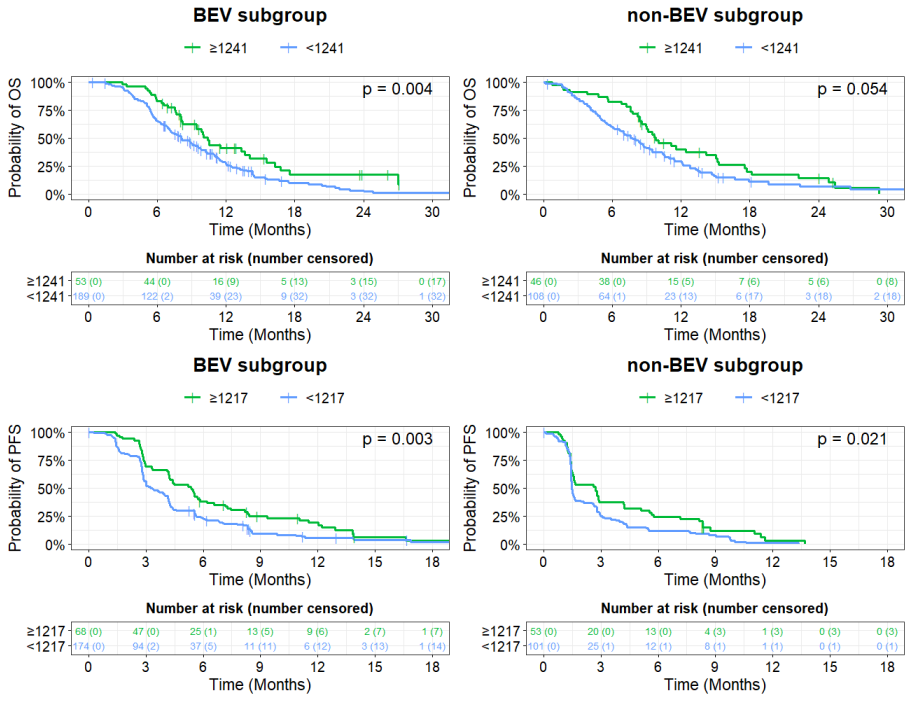
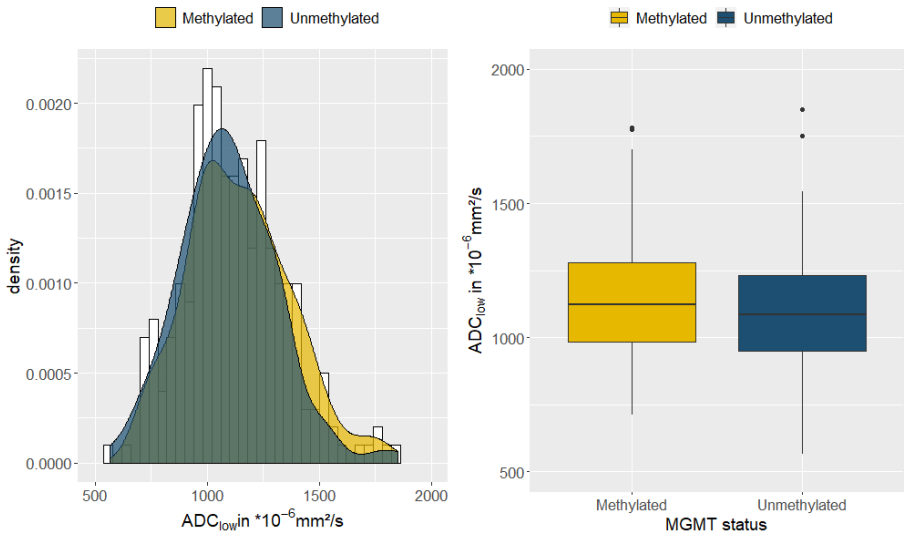


Figure 6. Relationship between ADC_{low} and MGMT promoter methylation status assessed with relative frequency histogram of ADC_{low} values according to MGMT status (left panel) and boxplots (right panel). Median ADC_{low} values were $1141*10^{-6}$ mm²/s for MGMT-methylated and $1095*10^{-6}$ mm²/s for MGMT-unmethylated tumors (p=0.13)



Data Supplement

Supplementary Table S1. Image acquisition parameters related to apparent diffusion coefficient (ADC) sequences as well as corresponding ADC histogram fits, comparing BEV and non-BEV subgroups.

Parameter	**BEV group** (n=242)	**non-BEV group** (n=154)	**p-value**
MRI vendor			
GE	101 (43)	57 (27)	0.27
Philips	51 (22)	29 (19)	
Siemens	84 (36)	66 (43)	
Toshiba	0 (0)	1 (1)	
Field strength			
1.5 T	67 (30)	52 (35)	0.50
3.0 T	49 (22)	31 (21)	
1.5 or 3 T	111 (49)	65 (44)	
ADC resolution			
In plane resolution (mm)			0.40
Median	0.94	0.98	
Interquartile Range (IQR)	0.94, 1.25	0.94, 1.25	
Slice thickness (mm)			0.14
Median	5.00	5.00	
Interquartile Range (IQR)	5.00, 5.00	5.00, 5.00	
B-values			
1000	71 (29)	61 (40)	0.09
1200	6 (2)	2 (1)	
Unknown	165 (68)	91 (59)	
Histogram fit			
Adjusted R^2			0.64
Median	0.99	0.99	
Interquartile Range (IQR)	0.98, 0.99	0.97, 0.99	
Width of bins (*10-6 mm^2/s)			1.00
Median	100	100	
Interquartile Range (IQR)	50, 100	50, 100	